Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2012, Article ID 296701, 13 pages
doi:10.1155/2012/296701

Research Article

D-Optimal Design for Parameter Estimation in Discrete-Time Nonlinear Dynamic Systems

Yu Liu,[1] Hamid Reza Karimi,[2] and Zhiwei Yu[1]

[1] *The Control Science and Engineering Department, Harbin Institute of Technology, Harbin 150001, China*
[2] *The Department of Engineering, Faculty of Engineering and Science, University of Agder, 4898 Grimstad, Norway*

Correspondence should be addressed to Yu Liu, liu_hit@126.com

Received 21 June 2012; Accepted 7 August 2012

Academic Editor: Bo Shen

An optimal input design method for parameter estimation in a discrete-time nonlinear system is presented in the paper to improve the observability and identification precision of model parameters. Determinant of the information matrix is used as the criterion function which is generally a nonconvex function about the input signals to be designed. To avoid the locally optimizing problem, a randomized design method is proposed by which a globally optimizing test plan other than input signals may be obtained. Then the randomized design can be approximated by a nonrandomized design about optimal inputs. An iterative algorithm integrated with dynamic programming is given and verified by a numerical example on experimental design for self-calibration tests of ISP system.

1. Introduction

Model parameters applied to computation or compensation in science and engineering, such as error model coefficients in INS (inertial navigation system), generally require much higher identification precision than in other applications. However, haphazard experiments not only lead to poor accuracy in parameter estimating, but also would make some parameters unobservable. A good experimental design can increase both the precision and the efficiency of a test [1] and then improve the precision of system identification or state estimation [2–5].

The field of system identification and filtering are relatively mature [6–10]; relevant experimental design methods have not made substantial advance yet. D-optimal design which allocates the experimental input variables by maximizing the determinant of information matrix of the system is recognized as the most effective method for an experimental design [11, 12]. For model parameter identification of a dynamic system, the

D-optimal design problem has a similar mathematical expression as the optimal control problem, but cannot be solved by the Pontryagin's maximum principle and dynamic programming method, due to particularity in the form of performance index [13].

On the other hand, even for a dynamic system with linear or low-order nonlinear model, the D-optimal design problems may involve global optimizing of nonconvex function and cannot achieve analytical solutions by traditional nonlinear programming methods. Although many numerical searching algorithms have been proposed to solve the nonconvex problem in global optimizing, such as genetic algorithm, simulated annealing algorithm, and so forth, most of them are either time-consuming or no guarantee of global optimization of searching results [14–16].

D-optimal design for randomized inputs is a convex optimization technique, in which the experimental variables are transformed to test plans. A test plan specifies different probability measures to each input variable in admissible set and one selects inputs for a particular trial of the experiment via randomization. The randomized design method is mainly used in regression design problems. Mehra introduces the method to optimal input design for parameter identification in a discrete-time MIMO linear system with process noise [11].

Morelli and Klein consider input design problem for LTI systems in aircraft flight tests, and the specific goals with test time optimization are achieved using principles of dynamic programming [17]. Neto et al. generalize the results to nonlinear dynamic systems and consider additive colored noises in measurement [18]. The cost function selected by Neto et al. is the trace of a dispersion matrix in which the autocorrelation matrix of the colored measurement noises is introduced. They solve the optimization problem by genetic algorithm.

Lintereur studies optimal trajectories for a 2-axis gimbaled test table by which errors in inertial systems caused by angular motion are calibrated [19]. The trace of covariance matrix computed by Kalman filtering is minimized using a conjugate gradient algorithm, but local minimum may be obtained.

In this paper, we propose a randomized design method for parameter estimation in discrete-time nonlinear dynamic systems with constraints on inputs. By this design method, the original nonconvex optimization problem can be solved by the convex optimization technique, and the global optimal maximum is guaranteed. An iterative algorithm is given and verified by a numerical example on experimental design for continuous tumbling self-calibration tests of ISP system.

2. Problem Statement

In this section we give a mathematical formulation of the D-optimal design problem, in which a time-varying MIMO nonlinear system with unknown model parameters is considered. For simplification in notation and deduction, the process noise is assumed to be zero here.

Consider the following nonlinear dynamic system:

$$\begin{aligned} x_{k+1} &= f(x_k, u_k, k, \theta), \quad x_0 = \overline{x}_0, \\ y_k &= h(x_k, t_k, \theta) + v_k, \quad k = 0, 1, \ldots, N, \end{aligned} \tag{2.1}$$

where $x_k = x(t_k)$ is a $n \times 1$ state vector, $\overline{x}_0$ is a constant vector, $u_k = u(t_k)$ is a $q \times 1$ input vector, $y_k = y(t_k)$ is a $p \times 1$ sampling output vector at moment k, and $v_k = v(t_k)$ is a $p \times 1$ measurement noise vector. v_k is the Gaussian white noise sequences with $E(v_k) = 0$ and $E(v_k v_\tau^T) = R_k \delta_{k,\tau}$, N is the number of output samples observed and is fixed. $\theta = [\theta_1\ \theta_2 \cdots \theta_m]^T$ denotes the $m \times 1$ vector of constant identifiable parameters, we estimate θ from the knowledge of $\{y_k, u_{k-1}, k = 1, \ldots, N\}$ and give an unbiased efficient estimator $\widehat{\theta}$ with covariance M^{-1}, where M is the Fisher information matrix. Therefore, the design problem is to select a series of inputs $u_k \in \Omega_u$ such that a suitable criterion function corresponding to the objectives of the identification experiment is optimized.

The Fisher information matrix is defined as follows:

$$M = \underset{\theta,Y}{E}\left[\left(\frac{\partial \log p(Y,\theta)}{\partial \theta}\right)\left(\frac{\partial \log p(Y,\theta)}{\partial \theta}\right)^T\right], \tag{2.2}$$

where Y denotes the set of observations $\{y_k,\ k = 1, \ldots, N\}$ and the expectation in (2.2) is taken over the sample space Ω_Y of observations and the parameter space Ω_θ of θ.

Using conditional expectations, M may be evaluated in two steps, first by computing $M'(\theta) = E_{Y|\theta}\{\bullet\}$ and then $M = E_\theta M'(\theta)$. The second step is generally more tedious and an *a priori* distribution $p(\theta)$ should be known exactly. Here a Taylor-series approximation is used to simplify the computation:

$$\begin{aligned} M'_{i,j}(\theta) = M'_{i,j}(\theta_0) &+ \left.\frac{\partial M'_{i,j}}{\partial \theta}\right|_{\theta_0}(\theta - \theta_0) \\ &+ \frac{1}{2}(\theta - \theta_0)^T \left.\frac{\partial^2 M'_{i,j}}{\partial \theta^2}\right|_{\theta_0}(\theta - \theta_0) + \cdots, \end{aligned} \tag{2.3}$$

where θ_0 is the *a priori* mean of $\theta, i, j = 1, \ldots, m$.

Retaining terms up to second order,

$$M_{i,j} = M'_{i,j}(\theta_0) + \frac{1}{2}\,\mathrm{tr}\left[\left.\frac{\partial^2 M'_{i,j}}{\partial \theta^2}\right|_{\theta_0} P_0\right], \tag{2.4}$$

where P_0 is the *a priori* covariance of θ.

The second term is typically small compared to the first term either because P_0 is small or $M'(\theta)$ is insensitive to θ.

The conditional likelihood function $L(\theta) = \log p(Y \mid \theta)$ for system (2.1) is given as follows:

$$L(\theta) = -\frac{Np}{2}\log(2\pi) - \frac{1}{2}\sum_{k=1}^{N}\left\{(v_k)^T R_k^{-1} v_k + \log|R_k|\right\}. \tag{2.5}$$

The matrix $M'(\theta)$ has elements

$$\begin{aligned} M'_{i,j}(\theta) &= E_Y\left(\frac{\partial L(\theta)}{\partial \theta_i}\frac{\partial L(\theta)}{\partial \theta_j}\right) \\ &= \sum_{K=1}^{N}\left\{\left(\frac{\partial h(x_k,t_k,\theta)}{\partial \theta_i}\right)^T R_k^{-1}\frac{\partial h(x_k,t_k,\theta)}{\partial \theta_j}\right\}. \end{aligned} \tag{2.6}$$

The sensitivity function is

$$\frac{\partial h(x_k,t_k,\theta)}{\partial \theta_i} = \frac{\partial h}{\partial x_k^T}x_{\theta_i,k} + \frac{\partial h}{\partial \theta_i}, \tag{2.7}$$

where $x_{\theta_i,k}$ is the partial derivative of $x(t)$ about θ_i at moment k, that is, $x_{\theta_i,k} = x_{\theta_i}(t_k)$ which meets the following equation:

$$x_{\theta_i,k+1} = \frac{\partial f}{\partial x_k^T}x_{\theta_i,k} + \frac{\partial f}{\partial \theta_i}, \quad x_{\theta_i,0} = 0, \ i = 1,\ldots,m. \tag{2.8}$$

Since $x_{\theta_i,k}$ and x_k all depend on the elements of U, where $U^{\mathrm{T}} = [u_0^T \cdots u_{N-1}^T] \in \Omega_U$ is the Nq-dimensional vector to be designed, we denote $M'(\theta)$ as $M'(\theta,U)$.

From (2.6), it is easy to get

$$M'(\theta,U) = \sum_{k=1}^{N}\left\{\left(\frac{\partial h(x_k,t_k,\theta)}{\partial \theta}\right)^T R_k^{-1}\frac{\partial h(x_k,t_k,\theta)}{\partial \theta}\right\}. \tag{2.9}$$

Also from (2.4), the Fisher information matrix is generally

$$M(U) = E_\theta M'(\theta,U) \approx M'(\theta_0,U). \tag{2.10}$$

There are many formulations of criterion function that measures the degree of observability about parameter θ, such as $\mathrm{tr}(M^{-1}(U))$ or $|M^{-1}(U)|$. A design which minimizes the scalar measure $|M^{-1}(U)|$ or maximizes $|M(U)|$ is called D-optimal, and it is equivalent to minimizing the volume of the uncertainty ellipsoid about parameter estimators. An important advantage of D-optimality is that it is invariant under scale changes in the parameters and linear transformations of the output.

Now, we choose $|M^{-1}(U)|$ as the criterion function and formulate the D-optimal design problem as follows:

$$\begin{aligned} &\min_{U\in\Omega_U}\left|M^{-1}(U)\right| \\ &\text{s.t.} \quad x_{k+1} = f(x_k,u_k,k,\theta), \quad x_0 = \overline{x}_0 \\ &\qquad x_{\theta_i,k+1} = \frac{\partial f}{\partial x_k^T}x_{\theta_i,k} + \frac{\partial f}{\partial \theta_i}, \quad x_{\theta_i,0} = 0, \ i = 1,\ldots,m. \end{aligned} \tag{2.11}$$

It should be pointed out that the problem in (2.11) cannot be solved by typical methods such as the Pontryagin's maximum principle and dynamic programming method since $|M^{-1}(U)|$ or $\mathrm{tr}(M^{-1}(U))$ cannot be transformed to the index form in multistage decision process. In fact, there also exists great difficulty in getting a numerical solution with global optimization since $|M^{-1}(U)|$ is not a convex function of U.

In the next section, we will present a randomized design method based on test plan theories.

3. D-Optimal Design Method

For a randomized input $U \in \Omega_U$ with probability measure $\xi(dU)$defined for all Borel sets and points of Ω_U, the definition of the information matrix is

$$M(\xi) = \int_{\Omega_U} M(U) \cdot \xi(dU), \tag{3.1}$$

where $\int_{\Omega_U} \xi(dU) = 1$.

If the probability measure is purely discrete, the information matrix is defined as follows:

$$M(\xi) = \sum_{i=1}^{l} \xi_i M(U_i), \tag{3.2}$$

where l is the number of spectrums, $\sum_{i=1}^{l} \xi_i = 1$, $0 \leq \xi_i \leq 1$.

$M(\xi)$ is linear in ξ, so the criteria $|M^{-1}(\xi)|$ or $\mathrm{tr}(M^{-1}(\xi))$ are convex functions of ξ, and optimization with respect to ξ gives globally optimizing design. However, we cannot find directly the optimal design ξ^* which minimizes $|M^{-1}(\xi)|$. Here, an iterative algorithm is proposed for searching ξ^* based on the following theorem [11].

Theorem 3.1. *Let ξ^* be the optimal design then the following are equivalent:*

(i) ξ^* *maximizes* $|M(\xi)|$,

(ii) ξ^* *minimizes* $\max_{U\in\Omega_U} \mathrm{tr}(M^{-1}(\xi)M(U))$,

(iii) $\max_{U\in\Omega_U} \mathrm{tr}(M^{-1}(\xi^*)M(U)) = \mathrm{tr}(M^{-1}(\xi^*)M(\xi^*)) = m$.

The D-optimal design ξ^* may be computed with the following algorithm.

Algorithm 3.2. *Step 1*. Start with any design ξ_0 such that $M(\xi_0)$ is nonsingular and let k=0.

Step 2. Compute $M(\xi_k)$ and $\mathrm{tr}(M^{-1}(\xi_k)M(U))$ using (2.1), (2.6)–(2.10).

Step 3. Maximize $\mathrm{tr}(M^{-1}(\xi_k)M(U))$ over $U \in \Omega_U$ and get U_k.

Step 4. If $\mathrm{tr}(M^{-1}(\xi_k)M(U_k)) = m$, stop. Otherwise, let $\xi_{k+1} = (1-\alpha_k)\xi_k + \alpha_k\xi(U_k)$, $0 < \alpha_k \leq 1$ where $\xi(U_k)$ is the design at the single point U_k.

Choose α_k such that $\sum_{k=0}^{\infty} \alpha_k = \infty$, $\lim_{k\to\infty}\alpha_k = 0$, $|M(\xi_{k+1})| \geq |M(\xi_k)|$.

Step 5. Set $k = k + 1$ and go to step 2.

The convergence of the above algorithm to the global maximum is proved in the appendix.

Remarks. (1) The optimal design can be depicted by the following set:

$$\{\xi_1, U_1; \xi_2, U_2; \ldots; \xi_l, U_l\}, \quad l \le m\frac{(m+1)}{2}. \tag{3.3}$$

This can be used in a manner of randomized strategies when the experiment can be repeated. If the experiment is to be conducted only once, a nonrandomized design involving only one input U should be preferable, that is, it assigns probability one to a particular U. Since the randomized design (3.3) has been derived, we can seek nonrandomized design U^* so that $|M(U^*)|$ approximates $|\sum_{i=1}^{l} \xi_i M(U_i)|$.

(2) Step 3 is most time-consuming computationally, and the criterion function $\operatorname{tr}(M^{-1}(\xi)M(U))$ is generally not a convex function of U. Only if model (2.1) can be reduced to a linear discrete-time system, it would be a quadratic functional of U. Using (2.9) and (2.10), we get

$$\begin{aligned} \operatorname{tr}\left(M^{-1}(\xi)M(U)\right) &\approx \operatorname{tr}\left(M^{-1}(\xi)M'(\theta_0, U)\right) \\ &= \sum_{k=1}^{N} \operatorname{tr}\left\{ M^{-1}(\xi)\left(\left.\frac{\partial h(x_k, t_k, \theta)}{\partial \theta}\right|_{\theta_0} \right)^T R_k^{-1} \left.\frac{\partial h(x_k, t_k, \theta)}{\partial \theta}\right|_{\theta_0} \right\}. \end{aligned} \tag{3.4}$$

Unlike the computing of determinant, the operations with trace and sum of matrix can exchange order. By (3.4), the optimization problem can be solved by using maximum principle or dynamic programming methods since the above equation possesses the form of criterion function in multistage decision process.

Therefore, by using randomization and Theorem 3.1 the solution to a highly nonlinear and nonconvex optimization problem, that is, minimization of $|M^{-1}(U)|$ is reduced to solving a relatively simpler optimization problem. This is mainly due to the fact that randomization produces convexity.

4. Simulation

In this section, we present a numerical example to verify the effectiveness of the proposed design method. The experiment to be designed is the continuous tumbling self-calibration test of an ISP system. Choose the determinant of information matrix as observability index and select the currents of gyro torquers or the command angular speed to the ISP as the experimental input variables, then the idea of D-optimal design can be applied to program the rotational trajectories of platform which represent the attitude and angular speed of platform at each moment.

First, we give the model equations of accelerometers and gyroscopes in the ISP system. The output equations of accelerometers that are also the observation equations are:

$$\mathbf{y} = \begin{bmatrix} 0 & A_z & -A_y \\ -A_z & 0 & A_x \\ A_y & -A_x & 0 \end{bmatrix} \begin{bmatrix} \psi_x \\ \psi_y \\ \psi_z \end{bmatrix} + \begin{bmatrix} k_{0x} + k_{1x}A_x \\ k_{0y} + k_{1y}A_y - \alpha_z A_x \\ k_{0z} + k_{1z}A_z + \alpha_y A_x - \alpha_x A_y \end{bmatrix} + \varepsilon, \tag{4.1}$$

where $\alpha_x, \alpha_y, \alpha_z$ represent the misalignment angles of accelerometers' input axes with respect to platform frame.

k_{0x}, k_{0y}, k_{0z} and k_{1x}, k_{1y}, k_{1z} represent the bias and scale factors of accelerometers.

y represents outputs of accelerometers; ε denotes the observation noises in outputs with zero mean and covariance matrix $\delta_y^2 I_{3\times 3}$ at each moment.

A_x, A_y, A_z are projections of gravitational acceleration on the ideal platform frame which is defined based on x accelerometer's input axis and initially aligned to north, west, vertical direction, unit g_0, where g_0 is the local gravitational acceleration value:

$$\begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} = \begin{bmatrix} \sin(a)\sin(c) - \cos(a)\sin(b)\cos(c) \\ \sin(a)\cos(c) + \cos(a)\sin(b)\sin(c) \\ \cos(a)\cos(b) \end{bmatrix}, \tag{4.2}$$

where a, b, c represent three ideal angular positions of platform gimbals from outer to inner, respectively, and meet the following differential equations:

$$\begin{aligned} \begin{bmatrix} \dot{a} \\ \dot{b} \\ \dot{c} \end{bmatrix} &= \begin{bmatrix} -\cos(a)\tan(b) & \cos(b) & \sin(a)\tan(b) \\ \sin(a) & 0 & \cos(a) \\ \cos(a)\sec(b) & 0 & -\sin(a)\sec(b) \end{bmatrix} \begin{bmatrix} \omega_{ec} \\ 0 \\ \omega_{es} \end{bmatrix} \\ &\quad - \begin{bmatrix} \cos(c)\sec(b) & -\sin(c)\sec(\mathrm{b}) & 0 \\ \sin(c) & \cos(c) & 0 \\ -\cos(c)\tan(b) & \sin(c)\tan(b) & -1 \end{bmatrix} \begin{bmatrix} tg_x \\ tg_y \\ tg_z \end{bmatrix}, \end{aligned} \tag{4.3}$$

where ω_{ec}, ω_{es} represent north and vertical components of rotational speed of the earth ω_e, respectively.

ψ_x, ψ_y, ψ_z in (4.1) represent the attitude errors between the practical platform frame and ideal one:

$$\begin{bmatrix} \dot{\psi}_x \\ \dot{\psi}_y \\ \dot{\psi}_z \end{bmatrix} = \begin{bmatrix} 0 & tg_z & -tg_y \\ -tg_z & 0 & tg_x \\ tg_y & -tg_x & 0 \end{bmatrix} \begin{bmatrix} \psi_x \\ \psi_y \\ \psi_z \end{bmatrix} + \begin{bmatrix} d_{0x} + d_{1x}tg_x - \gamma_{xz}tg_y + \gamma_{xy}tg_z \\ d_{0y} + d_{1y}tg_y + \gamma_{yz}tg_x - \gamma_{yx}tg_z \\ d_{0z} + d_{1z}tg_z - \gamma_{zy}tg_x + \gamma_{zx}tg_y \end{bmatrix}, \tag{4.4}$$

where $\gamma_{xy}, \gamma_{zy}, \gamma_{zx}, \gamma_{yx}, \gamma_{yz}, \gamma_{xz}$ represent the misalignment angles of gyroscopes' input axes with respect to platform frame.

d_{0x}, d_{0y}, d_{0z} and d_{1x}, d_{1y}, d_{1z} represent the fixed drifts and scale factor errors of gyro torquers.

$u(t) = [\, tg_x\ tg_y\ tg_z \,]^T$ represent equivalent command angular speed to the ISP in an ideal platform frame and are the input variables to be designed.

Choose ψ_x, ψ_y, ψ_z and a, b, c as the state vector $x(t)$; note that only the state of (4.4) is related to partial elements of unknown parameter vector θ, where

$$\begin{aligned} \theta^T &= \left[\theta_a^T\ \theta_g^T\right], \qquad \theta_a = \left[\alpha_x, \alpha_y, \alpha_z, k_{0x}, k_{0y}, k_{0z}, k_{1x}, k_{1y}, k_{1z}\right]^T, \\ \theta_g &= \left[\gamma_{xy}, \gamma_{zy}, \gamma_{zx}, \gamma_{yx}, \gamma_{yz}, \gamma_{xz}, d_{0x}, d_{0y}, d_{0z}, d_{1x}, d_{1y}, d_{1z}\right]^T. \end{aligned} \tag{4.5}$$

Therefore, the equations for $x_{\theta_i}(t)$ can be derived by making partial derivative of (4.4) about θ_g.

Rewrite (4.1), (4.3), and (4.4) for abbreviation as follows:

$$\begin{aligned} &\dot{x}_1(t) = A_1(u)x_1(t) + B_1(u)\theta_g, \qquad x_1(t_0) = x_{10}, \\ &\dot{x}_2(t) = f(x_2, u), \qquad x_2(t_0) = 0, \\ &y_k = H_1(x_{2,k})x_{1,k} + H_2(x_{2,k})\theta_a + \varepsilon_k, \quad k = 0, 1, \ldots, N, \end{aligned} \tag{4.6}$$

where $x_1(t) = [\varphi_x, \varphi_y, \varphi_z]^T$, $x_2(t) = [a, b, c]^T$.

Then the equations for $x_{\theta_i}(t)$ are as follows:

$$\dot{x}_{\theta_i}(t) = A_1(u)x_{\theta_i}(t) + B_{1,i}(u), \quad x_{\theta_i}(t_0) = 0, \; i = 1, \ldots, s, \tag{4.7}$$

where $B_{1,i}(u)$ denotes the ith column of matrix $B_1(u)$, and s is the dimension of θ_g.

$$M'(\theta, U) = \delta_y^{-2} \sum_{k=1}^{N} \begin{bmatrix} H_2^T(x_{2,k})H_2(x_{2,k}) & H_2^T(x_{2,k})H_1(x_{2,k})X_\theta \\ X_\theta^T H_1^T(x_{2,k})H_2(x_{2,k}) & X_\theta^T H_1^T(x_{2,k})H_1(x_{2,k})X_\theta \end{bmatrix} \tag{4.8}$$

Where $X_\theta = [x_{\theta_1,k}, \ldots, x_{\theta_s,k}]$.

It shows that $M'(\theta, U)$ is insensitive to θ since in (4.6) and (4.7), $x_{2,k}$ as well as $x_{\theta_i,k}$ is independent of the unknown parameter θ.

Thus, in this problem

$$M(U) = M'(\theta, U), \qquad M(\xi) = \sum_{i=1}^{l} \xi_i M(U_i) \tag{4.9}$$

for any θ and state equation for $x_1(t)$ can be deleted from the constraint conditions in (2.11).

Now we reformulate the design problem for self-calibration tests of ISP system as follows:

$$\begin{aligned} &\max_{U \in \Omega_U} |M(\xi)| \\ &\text{s.t.} \quad \dot{x}_2(t) = f(x_2, u), \qquad x_2(t_0) = 0, \\ &\qquad \dot{x}_{\theta_i}(t) = A_1(u)x_{\theta_i}(t) + B_{1,i}(u), \quad x_{\theta_i}(t_0) = 0, \; i = 1, \ldots, s, \end{aligned} \tag{4.10}$$

where $\Omega_U = \{U : u^- \le u_k \le u^+, \; k = 0, \ldots, N-1\}$ is the amplitude constraint on the precise command angular speed.

Since three state equations should be added to the constraint equations in (4.10) if one parameter in θ_g is to be estimated, the dynamic programming algorithm in Step 3 will be very time-consuming. Therefore, only 9 parameters in θ_a are considered in this example.

The initial design ξ_0 is chosen from a single-point design which maximizes $|M(U)|$ via a rough search by confining each element of the input signals $u(t)$ to a four-segment

Table 1: Design results by the propose algorithm.

i	1	2	3	4
$\log\lvert M(U_i)\rvert$	4.7	4.5	4.7	5.2
ξ_i	0.1	0.2	0.2	0.5

Table 2: Standard errors of parameter estimators ($\delta_y = 1$).

	α_x	α_y	α_z	k_{0x}	k_{0y}	k_{0z}	k_{1x}	k_{1y}	k_{1z}
U_1	0.67	0.77	0.76	0.4	0.43	0.43	0.7	0.67	0.63
U_4	0.65	0.57	0.53	0.36	0.36	0.38	0.53	0.56	0.66

square wave varying between $-10\,\omega_e$, 0, and $10\,\omega_e$. Figure 1 shows the convergence of $\mathrm{tr}(M^{-1}(\xi_k)M(U_k))$ after eleven times of iterations, and $|M(\xi_{k+1})|/|M(\xi_k)|$ tends to one which means $|M(\xi_k)|$ converges to a local maximum $|M(\xi^*)|$. Since $|M(\xi)|$ is a concave function about ξ, ξ^* is global optimizing as well.

The values of $\log|M(U_i)|$ at each supporting point U_i and measure ξ_i are listed in Table 1 with $\delta_y^2 = 1$. In addition, $\log|M(\xi^*)| = \log|\sum_i M(U_i)\xi_i| = 5.5$ and $\log|M(\xi(U'))| = \log|M(\sum_i U_i\xi_i)| = 1.7$. It shows that the randomized design ξ^* has better performance than any single-point design $\xi(U_i)$, whereas the linear combination $U' = \sum_i U_i\xi_i$, although it is also an admissible control input in Ω_U, shows poor performance and cannot be used as a nonrandomized approximation to the optimal test plan ξ^* here.

In Table 1 the 4th supporting point U_4 has the maximum $\log|M(U_i)|$ which approximates mostly to $\log|M(\xi^*)|$. So single-point design $\xi(U_4)$ is chosen as the nonrandomized approximation to ξ^*. Comparison of estimation errors between tests with U_4 and U_1 (the initial rough design) is in Table 2, which shows some improvement in budgets of estimation precision. The control input curves of U_4 are plotted in Figure 2 with unit ω_e. The whole test time is three hours, and the sampling time is 22.5 minutes. None of tg_x, tg_y, and tg_z in U_4 is the bang-bang type mainly because a constraint on angle $b(-\pi/6 \le b \le \pi/6)$ is also assumed. The projection of gravitational acceleration in the ideal platform frame is plotted in Figure 3.

The profiles of A_x, A_y, and A_z show that input axis of each accelerometer completes nearly a whole tumble in gravitational field which can provide sufficient stimulations to the error terms of accelerometers in practical testing.

5. Conclusion

A D-optimal design method for parameter estimation in nonlinear dynamic systems is presented based on test plan design theories. The corresponding iterative algorithm is proposed with a dynamic programming algorithm imbedded. The proof for the convergence of the algorithm is given as well. Simulation results on an optimal trajectory design problem in self-calibration test of ISP system demonstrate the effectiveness of the proposed algorithm.

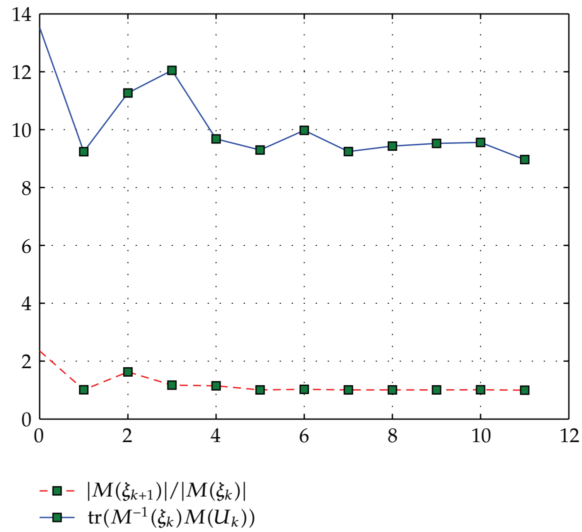

Figure 1: Convergence of the proposed algorithm.

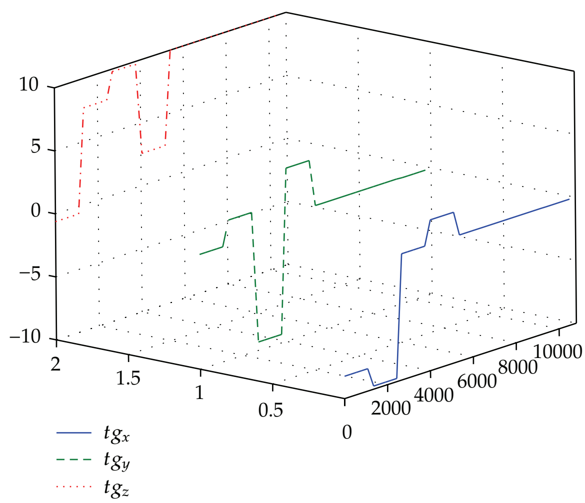

Figure 2: Curves of the supporting point U_4.

Appendix

Proof of the Convergence of Algorithm 3.2

Proof. In the algorithm, α_k is chosen such that

$$|M(\xi_0)| \leq |M(\xi_1)| \leq \cdots \leq |M(\xi^*)|. \tag{A.1}$$

Since any bounded monotone nondecreasing sequence converges, the sequence $|M(\xi_0)|, |M(\xi_1)|, \ldots, |M(\xi_k)|$ converges to some limit $|M(\widehat{\xi})|$.

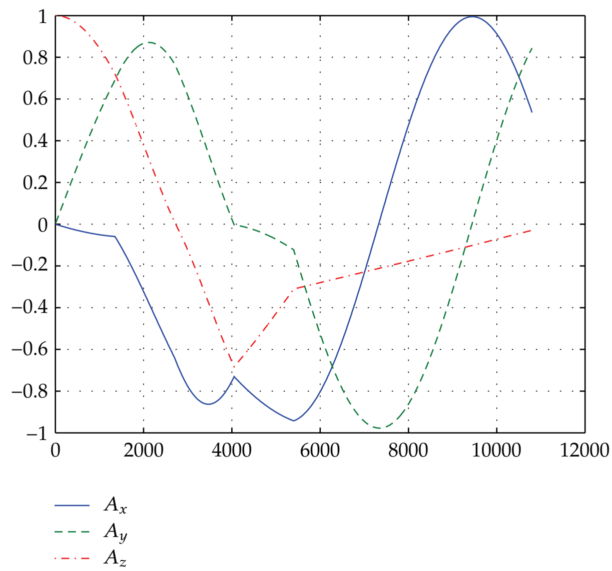

Figure 3: Acceleration projection in the ideal platform frame.

To prove $|M(\hat{\xi})| = |M(\xi^*)|$, we assume the contrary $|M(\hat{\xi})| < |M(\xi^*)|$. Then, by Theorem 3.1, for any k there is a constant η such that

$$\operatorname{tr}\left(M^{-1}(\xi_k)M(U_k)\right) - m > \eta > 0. \tag{A.2}$$

It follows that

$$\frac{\partial}{\partial \alpha_k}\log|M(\xi_{k+1})|\bigg|_{\alpha_k=0} = \operatorname{tr}\left[M^{-1}(\xi_k)(-M(\xi_k) + M(U_k))\right] > \eta. \tag{A.3}$$

For the smoothness of the function $\log|M(\xi_{k+1})|$ about α_k and by the assumption $\lim_{k\to\infty}\alpha_k = 0$, there is a positive integer s such that for any $k \geq s$, $(\partial/\partial\alpha_k)\log|M(\xi_{k+1})| \geq \eta$.

Now integrating both sides of the above inequality over α_k from 0 to α_k, one obtains

$$\frac{|M(\xi_{k+1})|}{|M(\xi_k)|} \geq \exp(\eta\alpha_k). \tag{A.4}$$

On the other hand, in view of the convergence of the sequence $|M(\xi_0)|, \ldots,$ for any small positive number γ, there is a positive integer n such that for any integer $p \geq q \geq n$, the following inequality holds:

$$|M(\xi_p)| - |M(\xi_q)| \leq \gamma. \tag{A.5}$$

Let $t = \max(n, s)$, then for any integer $p \geq q \geq t$,

$$\gamma \geq |M(\xi_p)| - |M(\xi_q)| \geq \left[\exp\left(\eta \sum_{k=q}^{p-1} \alpha_k\right) - 1\right] \cdot |M(\xi_q)|. \tag{A.6}$$

This means

$$\sum_{k=q}^{p-1} \alpha_k \leq \frac{1}{\eta}\left[\log(\gamma + |M(\xi_q)|) - \log|M(\xi_q)|\right], \tag{A.7}$$

which contradicts the assumption $\sum_{k=0}^{\infty} \alpha_k = \infty$.

Therefore, $|M(\hat{\xi})| = |M(\xi^*)|$, the global maximum is obtained by the algorithm. □

Acknowledgment

This work was supported by the Fundamental Research Funds for the Central Universities (Grant no. HIT. NSRIF. 2013037).

References

[1] G. C. Goodwin and R. L. Payne, *Dynamic system identification*, vol. 136, Academic Press, New York, NY, USA, 1977.
[2] B. Shen, Z. Wang, H. Shu, and G. Wei, "H_∞ filtering for nonlinear discrete-time stochastic systems with randomly varying sensor delays," *Automatica*, vol. 45, no. 4, pp. 1032–1037, 2009.
[3] B. Shen, Z. Wang, and Y. S. Hung, "Distributed H_∞-consensus filtering in sensor networks with multiple missing measurements: the finite-horizon case," *Automatica*, vol. 46, no. 10, pp. 1682–1688, 2010.
[4] D. Hongli, W. Zidong, and G. Huijun, "Distributed filtering for a class of time-varying systems over sensor networks with quantization errors and successive packet dropouts," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 3164–3173, 2012.
[5] H. Dong, Z. Wang, and H. Gao, "Robust H_∞ filtering for a class of nonlinear networked systems with multiple stochastic communication delays and packet dropouts," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 1957–1966, 2010.
[6] H. Jun, W. Zidong, G. Huijun et al., "Probability-guaranteed H-infinity finite-horizon filtering for a class of nonlinear time-varying systems with sensor saturation," *System & Control Letters*, vol. 61, no. 4, pp. 477–484, 2012.
[7] H. Jun, W. Zidong, G. Huijun et al., "Robust sliding mode control for discrete stochastic systems with mixed time-delays, randomly occurring uncertainties and randomly occurring nonlinearities," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 7, pp. 3008–3015, 2012.
[8] W. Ligang, S. Xiaojie, S. Peng et al., "Model approximation for discrete-time state-delay systems in the T-S fuzzy framework," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 2, pp. 366–378, 2011.
[9] Q. Jianbin, F. Gang, and Y. Jie, "A novel approach to filter design for T-S fuzzy discrete-time systems with time-varying delay," *IEEE Transactions on Fuzzy Systems*, vol. 17, no. 5, pp. 1044–1058, 2009.
[10] Y. Rongni, G. Huijun, and S. Peng, "Novel robust stability criteria for stochastic Hopfield neural networks with time delays," *IEEE Trans on Systems, Man and Cybernetics B*, vol. 39, no. 11, pp. 467–474, 2009.
[11] R. K. Mehra, "Optimal input signals for parameter estimation in dynamic systems—survey and new results," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 753–768, 1974.
[12] V. V. Fedorov, *Theory of Optimal Experiments*, Academic Press, New York, NY, USA, 1972.
[13] A. Lew and H. Mauch, *Dynamic Programming, A Computational Tool*, Springer, 2007.

[14] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Boston, Mass, USA, 1989.

[15] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[16] A. A. Zhigljavsky, *Theory of Global Random Search*, vol. 65, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.

[17] E. A. Morelli and V. Klein, "Optimal input design for aircraft parameter estimation using dynamic programming principles," in *Proceedings of the AIAA Atmospheric Flight Mechanics Conference*, pp. 249–259, 1990.

[18] N. Neto, E. Hemerly, and L. C. S. Goes, "Aircraft parameter estimation experiment design considering measurement colored residuals," in *Proceedings of the AIAA Atmospheric Flight Mechanics Conference*, pp. 1–12, 1990.

[19] L. J. Lintereur, *Optimal test trajectories for calibrating inertial systems [M.S. thesis]*, Massachusetts Institute of Technology, 1996.